

Quality Testing Standards – A Starter Kit for States

Version 6.17.2020

What is the QTS?

New Meridian Corporation has developed the *Quality Testing Standards and Criteria for Comparability Claims* (QTS) to provide guidance to states that are interested in including New Meridian content and would like to either keep reporting scores on the New Meridian Scale or use the New Meridian performance levels; that is, the state wishes to make comparability claims related to the New Meridian Scale or performance levels.

To support the state's desired comparability claims and provide for processes such as federal peer review, New Meridian recommends that each interested state collect and submit evidence demonstrating that these types of comparisons are technically defensible. This evidence may be evaluated by independent expert reviewers to determine if the desired comparisons can be supported. If the desired comparisons cannot be supported, the reviewers should provide constructive and actionable feedback on what the state needs to do to support the comparability claims. This is referred to as the *QTS comparability review process*.

This document is intended for anyone supporting a state that is considering New Meridian content for its assessment program. It includes a [questionnaire](#) for collecting information about the state's goals and priorities and a [checklist](#) of potential evidence that the state can provide for the QTS comparability review process. The document is part of a set of materials¹ that together define a system for evaluating the types of comparability claims that can be made by states using New Meridian content. If you have any questions about this document or the QTS process, please email info@newmeridiancorp.org.

Questionnaire about the State's Assessment Plan

The purpose of this questionnaire is to collect high-level information about the state's plan for its assessment that includes New Meridian content. We understand that some aspects of the state's assessment program may be still be under discussion or undergoing changes, so please answer the following questions based on what the state's current plans are for the *upcoming operational administration* of the assessment with New Meridian content.

1. **Comparability Claims.** What type of comparability claims does the state intend to make with the assessments that include New Meridian content?

¹ Additional materials include the *Quality Testing Standards and Criteria for Comparability Claims* (QTS main document), *Standard Processes*, and *Comparability Review Guidelines*.

✓	Type of Comparability Claim	Working Definition (per the QTS)
	Scale Score Comparability	<i>If a student taking the state's assessment with New Meridian content took one of the test forms offered by New Meridian, would he or she obtain the same scale score?</i>
	Readiness Comparability	<i>If a student taking the state's assessment with New Meridian content took one of the test forms offered by New Meridian, would he or she receive the same designation in terms of college and career readiness²?</i>
	No Comparability	<i>The state is only interested in licensing content in the New Meridian item bank and is not interested in making any comparability claims to other assessment with New Meridian content.</i>

If the state's intent is **No Comparability**, then STOP after Question 1. There is no need to answer the other questions.

2. **Use Case.** Which of the following best describes how the state's plan for its assessments with New Meridian content?

✓	Use Case	Description
	State-licensed New Meridian forms	<i>The state licenses New Meridian assessment content with test forms (flagship or ABO) designed to match the specifications and blueprints for New Meridian test forms. The state contracts its own vendor for the other steps in the operational administration process, including delivery, scoring and reporting.</i>
	State-licensed New Meridian forms, supplemented with state-developed content	<i>The state licenses New Meridian assessment content, but also includes content from its own (state-specific) item bank. The test forms are designed to match the specifications and blueprints for the New Meridian test forms. The state contracts its own vendor for the other steps in the operational administration process, including delivery, scoring and reporting.</i>

² "College and career readiness" in this context refers to both being *on track* for college and careers (i.e., grade-level readiness in elementary and middle school) and *ready* for college and careers (in high school).

v	Use Case	Description
	State developed assessments, supplemented with New Meridian content	<i>The state develops its own test items but also licenses New Meridian assessment content. The test forms are designed to match state-developed test specifications and blueprints. The state contracts its own vendor for test development, administration, scoring and reporting.</i>

Alternative use case (please describe):

3. **Grade and Content Area.** For which grade levels and content area tests is the state planning to include content from the New Meridian item bank?

Grade level and content area	Yes	No	Not Sure
Grade 3 English Language Arts/Literacy			
Grade 4 English Language Arts/Literacy			
Grade 5 English Language Arts/Literacy			
Grade 6 English Language Arts/Literacy			
Grade 7 English Language Arts/Literacy			
Grade 8 English Language Arts/Literacy			
Grade 9 English Language Arts/Literacy			
Grade 10 English Language Arts/Literacy			
Grade 11 English Language Arts/Literacy			

Grade level and content area	Yes	No	Not Sure
Grade 3 Mathematics			
Grade 4 Mathematics			
Grade 5 Mathematics			
Grade 6 Mathematics			
Grade 7 Mathematics			
Grade 8 Mathematics			
Algebra 1			
Algebra 2			
Geometry			
Integrated Mathematics 1			
Integrated Mathematics 2			
Integrated Mathematics 3			

4. **Assessment Purpose.** For which of the following purposes does the state plan use the assessments with New Meridian content?

Purpose	Yes	No	Not Sure
Summative (evaluate student achievement at the end of the school year)			
Interim/Formative (evaluate students and inform teacher instruction during the school year)			
Predictive (predict student performance on summative assessment at the end of the school year)			
Measure annual student growth or progress			
Teacher evaluation			
Grade promotion			
High school graduation			
School accountability ratings			
Other purposes (please specify):			

5. **Within-State Comparisons.** Which of the following comparisons does the state plan to make with the results from its assessment with New Meridian content?

Comparison within the State	Yes	No	Not Sure
Student performance across years (i.e., growth)			
Aggregated (e.g., school or district) performance within and across years			
Other within-state comparisons (please specify):			

6. **Cross-State Comparisons.** Which of the following comparisons does the state plan to make with other states with the results from its assessment with New Meridian content?

Comparison with Other States	Yes	No	Not Sure
Average scale scores			
Percentage of students in each performance level			
Percentage of students ready for the next course, grade level or for college and careers (i.e., attain Level 4 or higher)			
Other cross-state comparisons (please specify):			

7. **Blueprint.** For its operational test forms, does the state plan to include only items from New Meridian bank or include a mixture of state-developed items and items from New Meridian's bank?
- ☐ Only items from New Meridian's bank
- ☐ Mixture of state-developed items and items from New Meridian's bank

If the state plans to include a mixture of state-developed and New Meridian items on its operational test forms,

- a. Approximately what percentage does it plan to include from each source? (Please provide a range if the % of items differ across grade levels and content areas)

Source	Approximate %	
	ELA	Mathematics
New Meridian bank		
State developed		

- b. Are the state-developed items designed to measure the Common Core State Standards (CCSS)?
- ☐ Yes
- ☐ No

If so, are the CCSS measured by the state-developed items, the [original standards](#), a slightly amended version of the standards, or a significantly modified version of the standards.

- ☐ Original CCSS
- ☐ Slightly amended CCSS
- ☐ Significantly modified CCSS

8. **Item Types or Features.** Which of the following item types from the New Meridian bank is the state planning to include on its assessments?³

Item Type	Yes	No	Not Sure
Selected response (multiple choice or multiple select)			
Constructed response (writing/prose constructed response tasks)			
Technology enhanced (e.g., drag and drop, hot spot, fill-in-the-blank, inline choice, equation editor etc.)			

³ Examples of technology-enhanced item types in the New Meridian bank can be found at <https://dc.mypearsonsupport.com/tutorial/>

9. **Accommodations.** Which of the following accommodations does the state plan to allow on its assessments with New Meridian content?⁴

Accommodation	Yes	No	Not Sure
Text-to-Speech (online only)			
American Sign Language (online only)			
Closed Captioning (online only)			
Large Print (paper only)			
Braille (paper only)			
Other accommodations (please specify):			

10. **Languages.** In which language(s) does the state plan to administer its assessment with New Meridian content?

Language	Yes	No	Not Sure
English			
Spanish (mathematics only)			
Other languages (please specify):			

Does the state plan to include native language directions⁵ for its assessments with New Meridian content? If so, please specify in which languages?

- ☐ No
- ☐ Yes, please specify the languages:

11. **Administration Mode.** In which mode(s) does the state plan to administer its assessments with New Meridian content?

Administration Mode	Yes	No	Not Sure
Paper			

⁴ Demonstration of the available online accommodation for the New Meridian test forms can be found at <https://dc.mypearsonsupport.com/tutorial/>

⁵ For the New Meridian test forms, native language directions are provided as translated scripts in Arabic, Mandarin Chinese, Haitian Creole, Navajo, Polish, Portuguese, Russian, Spanish, Urdu, and Vietnamese.

Computer (desktop or laptop)			
Tablet			
Other administration modes (please specify):			

If the state plans to administer its assessments online (e.g., on computer or tablet), are test forms fixed forms or adaptive?

- ☐ Fixed forms
☐ Adaptive (either multistage or item adaptive)
☐ Not sure

12. **Testing Time.** If the state plans to include time limits for its assessments with New Meridian content, what are the planned time limits? (Please provide a range if the time limits differ across grades within any of the grade spans)

Content Area	Grades 3-5	Grades 6-8	High School
ELA			
Mathematics			

13. **Testing Window.** For how long and at what time of year does the state plan to have its testing windows for the assessments with New Meridian content? (Please fill out the rows that apply to the state's testing program.)

Elementary and Middle School (Grades 3-8)

Administration	Length	Start	End
Spring			
Online Testing			
Paper Testing			
Fall			
Online Testing			
Paper Testing			
Summer			
Online Testing			
Paper Testing			

High School

Administration	Length	Start	End
Spring			
Online Testing			
Paper Testing			
Fall			
Online Testing			

Administration	Length	Start	End
<i>Paper Testing</i>			
Summer			
<i>Online Testing</i>			
<i>Paper Testing</i>			

14. **Performance Scoring.** If the state plans to include constructed-response items (e.g., writing tasks) from the New Meridian bank on its assessments, how does it plan to score them?
- ☐ Only human scoring
 - ☐ Combination of human and automated/artificial intelligence (AI) scoring
 - ☐ Only AI scoring
 - ☐ Other (please specify): _____
15. **Item Parameters.** Does the state plan to use the existing IRT parameters (on the New Meridian scale) for the items from the New Meridian bank?
- ☐ Yes
 - ☐ No
 - ☐ Not sure
16. **Standard Setting.** How does the state plan to establish performance level descriptors (PLDs) and cut scores on its assessments with New Meridian content?
- ☐ Adopt the New Meridian PLDs and cut scores for all performance levels (Levels 1 to 5)
 - ☐ Only adopt the New Meridian PLDs and cut score that indicates readiness for the next course, grade level or for college and careers (Level 4) only
 - ☐ Only adopt the New Meridian PLDs, but set states-specific cut scores
 - ☐ Develop state-specific PLDs and set states-specific cut scores
17. **Score Reporting.** Which of the following New Meridian reporting scales does the state plan to use to report results for its assessments with New Meridian content?⁶

New Meridian Reporting Scale	Yes	No	Not Sure
Scale for summative scores			
Scale for claim (reading and writing) scores			
Subclaim classifications			

⁶ Details about the New Meridian reporting scales can be found at <https://www.isbe.net/Documents/IAR-Score-Interpretation-Guide.pdf>.

Comparability Evaluation Checklist

This section provides suggested lists of evidence that a state can consider collecting to support its intended comparability claims. The checklists are based on the supporting evidence described in New Meridian’s *Quality Testing Standards and Criteria for Comparability Claims* (QTS) and follows the same organizational structure of the six key aspects of a testing program: item and test development, fairness and accessibility, test administration, item scoring, psychometrics, and standard setting. If the state is submitting evidence for the QTS comparability review process, consider filling out the table for each aspect of the state’s assessment program with the following information:

- Checkmark (v):** If the evidence is (or will be) available, please check this column. If this evidence does not apply to the state’s assessment program, please leave this column and the remaining columns blank. For example, if the test forms for the state’s assessment *only* includes New Meridian content, then it is not necessary to provide evidence under *Item and Test Development*. If the state is *only* using items from the New Meridian bank and the accessibility features (such as online accommodations and language translations) that come with the New Meridian content, then it is not necessary to provide evidence under *Fairness and Accessibility*.
- Source:** Please provide information about the name and location of the evidence. If the evidence is available on a publicly accessible web site, please provide its URL. If the evidence will be emailed to the reviewers, please indicate “Send via email” and include the name of the documents or materials. If the evidence includes confidential data, please indicate “Secure location” and contact the reviewers to set up secure access to the evidence.
- Notes:** Please include additional information about the evidence that will help in the evaluation process.

To the extent practicable, please provide evidence for the state’s assessments with New Meridian content from the most current administration. We understand that some of the required evidence may not be available at this time. In such cases, please provide an approximate timeframe (in the *Source* column) for when such documentation or materials will be available. Evidence from the previous administration of the state’s assessments with New Meridian content can also be provided in lieu of that from the most current administration.

Item and Test Development

✓	Supporting Evidence for...	Source	Notes
	Test purpose, target population, and intended uses		
	Assessed content standards, item types, rubrics, blueprints, test formats, eligible content, and time limits, along with the rationale for the test design decisions		
	Procedures for review of test items by subject matter experts		
	Field testing and data review procedures		
	Forms construction and review procedures		

Example Sources of Evidence for Item and Test Development

- Documentation or web pages about the testing program and its assessments
- Documentation or web pages about the assessed curriculum (state-mandated or district-selected), content standards, and claims structure
- Item development specifications and processes, and qualitative and quantitative item review and piloting procedures
- Test development and review procedures, including test blueprints or specifications
- Forms or test construction specifications, including test construction targets, and forms review and approval procedures
- Materials or minutes for educator or stakeholder committee meetings
- Content alignment study reports

Fairness and Accessibility

✓	Supporting Evidence for...	Source	Notes
	Universal design principles		
	Accommodations for English learners and students with disabilities		
	Procedures used to translate forms for students for whom English is a second language		

Example Sources of Evidence for Fairness and Accessibility

- Accommodations manuals, tutorials or guides
- Test translation or trans-adaptation guidelines
- Materials or minutes from bias and sensitivity review committee meetings
- Evidence supporting the fairness of assessment results for all students and disaggregated student groups
- Research reports related to accessibility, universal design principles, and the validity of accommodations and language translations
- Annual technical reports or manuals

Test Administration

✓	Supporting Evidence for...	Source	Notes
	Training and instructions provided to test administrators and coordinators		
	Instructions given to test takers		
	Information about the modes of administration		
	Details about test security protocols		
	Evidence that supports accessibility of the test to all students as part of the test administration		

Example Sources of Evidence for Test Administration

- District and/or campus test administrator and coordinator manuals and training materials
- Practice tests or tutorials for test takers
- Documentation or web pages about online testing interface for computer-based tests
- Specification about technology requirements for computer-based tests
- Research reports on administration mode (paper vs. computer) or device (computer vs. tablet) comparability studies
- Specification about adaptive testing procedures or process
- Test security and administration procedures
- Documentation about data forensics analyses
- Accommodations manuals, tutorials or guides

Item Scoring

✓	Supporting Evidence for...	Source	Notes
	Training and qualification procedures for human scorers		
	Protocols for both machine and human scoring processes		
	Evidence that the scoring process is fair to all students		
	If used, validation of automated scoring processes		

Example Sources of Evidence for Item Scoring

- Documentation about machine scoring rules, test maps, test deck, and quality assurance procedures
- Documentation about recruitment and qualification of human scorers
- Training materials for human scorers
- Procedures for calibrating scoring throughout the human scoring process
- Procedures and criteria for monitoring human scorer quality
- Sample scoring materials, including rubric and anchor, training, qualifying, and validity sets
- Reports about the human scoring process, including inter-rater reliability, score point distribution, and validity sets results
- Research reports about the validity of automated/artificial intelligence (AI) scoring and fairness of the scoring process to all students
- Annual technical reports or manuals

Psychometrics

✓	Supporting Evidence for...	Source	Notes
	Choice of psychometric models		
	Scaling and equating design and procedures, including quality control processes		
	Analysis of disaggregated student groups		
	Sampling, including purpose and methodology		
	Other psychometric procedures or analyses		

Example Sources of Evidence for Psychometrics

- Operational psychometrics procedures specifications or guidelines
- Specification about adaptive testing methodology (routing logic, stopping rules, content balancing and exposure control criteria, etc.)
- Equating and scaling specifications, including quality assurance procedures and criteria
- Documentation about the choice of measurement model, how scales were established, and scale score characteristics
- Documentation about sampling for scaling, equating, or other psychometric analyses
- Procedures and results of any analysis of disaggregated student group performance on operational items
- Analysis or studies that support the reliability and validity of test scores
- Research plans or reports that support the comparability of test scores between the state's assessment and New Meridian
- Annual technical reports or manuals

Standard Setting

✓	Supporting Evidence for...	Source	Notes
	Achievement or performance level descriptors (ALDs or PLDs)		
	Standard setting methodology and procedures		
	Empirical support for the cut scores		

Example Sources of Evidence for Standard Setting

- Procedures for establishing policy descriptors, and ALDs or PLDs
- Research studies or documentation that supports the standard setting methodology and procedures
- Standard setting specifications that include details about each step of the standard setting process
- Description of all stakeholders involved in the standard setting process
- Standard setting meeting materials, including agenda, facilitator slides, panelist forms, and example of feedback data
- Studies or empirical data that support the validity of cut scores across grade levels and/or content areas
- External validity research studies, such as correlational, linking, and benchmarking studies
- Research reports on consequential validity
- Standard setting technical report or summary